```
# Uninstall the potentially incompatible versions
!pip uninstall -y gym numpy tensorflow jax

# Install a compatible version of gym and a specific version of numpy
!pip install gym==0.26.2 numpy==1.23.5 tensorflow jax
```

```
Found existing installation: gym 0.26.2
Uninstalling gym-0.26.2:
  Successfully uninstalled gym-0.26.2
Found existing installation: numpy 2.1.3
Uninstalling numpy-2.1.3:
  Successfully uninstalled numpy-2.1.3
Found existing installation: tensorflow 2.19.0
Uninstalling tensorflow-2.19.0:
  Successfully uninstalled tensorflow-2.19.0
Found existing installation: jax 0.6.2
Uninstalling jax-0.6.2:
  Successfully uninstalled jax-0.6.2
Collecting gym==0.26.2
  Using cached gym-0.26.2-py3-none-any.whl
Collecting numpy==1.23.5
  Using cached numpy-1.23.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.3 kB)
Collecting tensorflow
  Using cached tensorflow-2.19.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (4.1 kB)
Collecting jax
  Using cached jax-0.6.2-py3-none-any.whl.metadata (13 kB)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.11/dist-packages (from gym==0.26.2) (3.1.1)
Requirement already satisfied: gym_notices>=0.0.4 in /usr/local/lib/python3.11/dist-packages (from gym==0.26.2) (0.0.8)
Requirement already satisfied: absl-py>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (1.4.0)
Requirement already satisfied: astunparse>=1.6.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (1.6.3)
Requirement already satisfied: flatbuffers>=24.3.25 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (25.2.10)
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (0.6.0)
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (0.2.0)
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (18.1.1)
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (3.4.0)
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (from tensorflow) (24.2)
Requirement already satisfied: protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<6.0.0dev,>=3.20.3 in /usr/local/lib/python3.11/dist-packages (from tensorflow)
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (2.32.3)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from tensorflow) (75.2.0)
Requirement already satisfied: six>=1.12.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (1.17.0)
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (3.1.0)
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (4.14.0)
Requirement already satisfied: wrapt>=1.11.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (1.17.2)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (1.73.0)
Requirement already satisfied: tensorboard~=2.19.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (2.19.0)
Requirement already satisfied: keras>=3.5.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (3.8.0)
INFO: pip is looking at multiple versions of tensorflow to determine which version is compatible with other requirements. This could take a while.
Collecting tensorflow
  Downloading tensorflow-2.18.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (4.1 kB)
Collecting tensorboard<2.19,>=2.18 (from tensorflow)
  Downloading tensorboard-2.18.0-py3-none-any.whl.metadata (1.6 kB)
Collecting tensorflow
  Downloading tensorflow-2.18.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (4.1 kB)
  Downloading tensorflow-2.17.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (4.2 kB)
Requirement already satisfied: h5py>=3.10.0 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (3.14.0)
Collecting ml-dtypes<0.5.0,>=0.3.1 (from tensorflow)
  Downloading ml_dtypes-0.4.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (20 kB)
Collecting protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3 (from tensorflow)
  Downloading protobuf-4.25.8-cp37-abi3-manylinux2014_x86_64.whl.metadata (541 bytes)
Collecting tensorboard<2.18,>=2.17 (from tensorflow)
  Downloading tensorboard-2.17.1-py3-none-any.whl.metadata (1.6 kB)
```

```
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.11/dist-packages (from tensorflow) (0.37.1)
Requirement already satisfied: jaxlib<=0.6.2,>=0.6.2 in /usr/local/lib/python3.11/dist-packages (from jax) (0.6.2)
INFO: pip is looking at multiple versions of jax to determine which version is compatible with other requirements. This could take a while.
Collecting jax
  Downloading jax-0.6.1-py3-none-any.whl.metadata (13 kB)
Collecting jaxlib<=0.6.1,>=0.6.1 (from jax)
  Downloading jaxlib-0.6.1-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.2 kB)
Collecting jax
  Downloading jax-0.6.0-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.6.0,>=0.6.0 (from jax)
  Downloading jaxlib-0.6.0-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.2 kB)
Collecting jax
  Downloading jax-0.5.3-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.5.3,>=0.5.3 (from jax)
  Downloading jaxlib-0.5.3-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.2 kB)
Collecting jax
  Downloading jax-0.5.2-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.5.2,>=0.5.1 (from jax)
  Downloading jaxlib-0.5.1-cp311-cp311-manylinux2014_x86_64.whl.metadata (978 bytes)
Collecting jax
  Downloading jax-0.5.1-py3-none-any.whl.metadata (22 kB)
  Downloading jax-0.5.0-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.5.0,>=0.5.0 (from jax)
  Downloading jaxlib-0.5.0-cp311-cp311-manylinux2014_x86_64.whl.metadata (978 bytes)
Collecting jax
  Downloading jax-0.4.38-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.38,>=0.4.38 (from jax)
  Downloading jaxlib-0.4.38-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.0 kB)
INFO: pip is still looking at multiple versions of jax to determine which version is compatible with other requirements. This could take a while.
Collecting jax
  Downloading jax-0.4.37-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.37,>=0.4.36 (from jax)
  Downloading jaxlib-0.4.36-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.0 kB)
Collecting jax
  Downloading jax-0.4.36-py3-none-any.whl.metadata (22 kB)
  Downloading jax-0.4.35-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.35,>=0.4.34 (from jax)
  Downloading jaxlib-0.4.35-cp311-cp311-manylinux2014_x86_64.whl.metadata (983 bytes)
Collecting jax
  Downloading jax-0.4.34-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.34,>=0.4.34 (from jax)
  Downloading jaxlib-0.4.34-cp311-cp311-manylinux2014_x86_64.whl.metadata (983 bytes)
Collecting jax
  Downloading jax-0.4.33-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.33,>=0.4.33 (from jax)
  Downloading jaxlib-0.4.33-cp311-cp311-manylinux2014_x86_64.whl.metadata (983 bytes)
INFO: This is taking longer than usual. You might need to provide the dependency resolver with stricter constraints to reduce runtime. See https://pip.pypa.io/warnings/backt
Collecting jax
  Downloading jax-0.4.31-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.31,>=0.4.30 (from jax)
  Downloading jaxlib-0.4.31-cp311-cp311-manylinux2014_x86_64.whl.metadata (983 bytes)
Collecting jax
  Downloading jax-0.4.30-py3-none-any.whl.metadata (22 kB)
Collecting jaxlib<=0.4.30,>=0.4.27 (from jax)
  Downloading jaxlib-0.4.30-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.0 kB)
```

```
Requirement already satisfied: scipy>=1.9 in /usr/local/lib/python3.11/dist-packages (from jax) (1.15.3)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from astunparse>=1.6.0->tensorflow) (0.45.1)
Requirement already satisfied: rich in /usr/local/lib/python3.11/dist-packages (from keras>=3.5.0->tensorflow) (13.9.4)
Requirement already satisfied: namex in /usr/local/lib/python3.11/dist-packages (from keras>=3.5.0->tensorflow) (0.1.0)
Requirement already satisfied: optree in /usr/local/lib/python3.11/dist-packages (from keras>=3.5.0->tensorflow) (0.16.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.21.0->tensorflow) (3.4.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.21.0->tensorflow) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.21.0->tensorflow) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.21.0->tensorflow) (2025.6.15)
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.11/dist-packages (from tensorboard<2.18,>=2.17->tensorflow) (3.8)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from tensorboard<2.18,>=2.17->tensorflow) (0.7.2)
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from tensorboard<2.18,>=2.17->tensorflow) (3.1.3)
Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.11/dist-packages (from werkzeug>=1.0.1->tensorboard<2.18,>=2.17->tensorflow) (3.0.2)
Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/dist-packages (from rich->keras>=3.5.0->tensorflow) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from rich->keras>=3.5.0->tensorflow) (2.19.1)
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.11/dist-packages (from markdown-it-py>=2.2.0->rich->keras>=3.5.0->tensorflow) (0.1.2)
Using cached numpy-1.23.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.1 MB)
Downloading tensorflow-2.17.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (601.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 601.3/601.3 MB 1.9 MB/s eta 0:00:00
Downloading jax-0.4.30-py3-none-any.whl (2.0 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 2.0/2.0 MB 72.0 MB/s eta 0:00:00
Downloading jaxlib-0.4.30-cp311-cp311-manylinux2014_x86_64.whl (79.6 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 79.6/79.6 MB 9.0 MB/s eta 0:00:00
Downloading ml_dtypes-0.4.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.2 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 2.2/2.2 MB 64.8 MB/s eta 0:00:00
Downloading protobuf-4.25.8-cp37-abi3-manylinux2014_x86_64.whl (294 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 294.9/294.9 kB 20.6 MB/s eta 0:00:00
Downloading tensorboard-2.17.1-py3-none-any.whl (5.5 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 5.5/5.5 MB 80.7 MB/s eta 0:00:00
Installing collected packages: protobuf, numpy, tensorboard, ml-dtypes, gym, jaxlib, tensorflow, jax
  Attempting uninstall: protobuf
    Found existing installation: protobuf 5.29.5
    Uninstalling protobuf-5.29.5:
      Successfully uninstalled protobuf-5.29.5
  Attempting uninstall: tensorboard
    Found existing installation: tensorboard 2.19.0
    Uninstalling tensorboard-2.19.0:
      Successfully uninstalled tensorboard-2.19.0
  Attempting uninstall: ml-dtypes
    Found existing installation: ml_dtypes 0.5.1
    Uninstalling ml_dtypes-0.5.1:
      Successfully uninstalled ml_dtypes-0.5.1
  Attempting uninstall: jaxlib
    Found existing installation: jaxlib 0.6.2
    Uninstalling jaxlib-0.6.2:
      Successfully uninstalled jaxlib-0.6.2
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts
pymc 5.23.0 requires numpy>=1.25.0, but you have numpy 1.23.5 which is incompatible.
dopamine-rl 4.1.2 requires gym<=0.25.2, but you have gym 0.26.2 which is incompatible.
albucore 0.0.24 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
grpcio-status 1.71.0 requires protobuf<6.0dev,>=5.26.1, but you have protobuf 4.25.8 which is incompatible.
xarray-einstats 0.9.0 requires numpy>=1.25, but you have numpy 1.23.5 which is incompatible.
albumentations 2.0.8 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
flax 0.10.6 requires jax>=0.5.1, but you have jax 0.4.30 which is incompatible.
chex 0.1.89 requires numpy>=1.24.1, but you have numpy 1.23.5 which is incompatible.
```

```
db-dtypes 1.4.3 requires numpy>=1.24.0, but you have numpy 1.23.5 which is incompatible.
tensorflow-decision-forests 1.11.0 requires tensorflow==2.18.0, but you have tensorflow 2.17.1 which is incompatible.
tensorflow-text 2.18.1 requires tensorflow<2.19,>=2.18.0, but you have tensorflow 2.17.1 which is incompatible.
orbax-checkpoint 0.11.15 requires jax>=0.5.0, but you have jax 0.4.30 which is incompatible.
bigframes 2.6.0 requires numpy>=1.24.0, but you have numpy 1.23.5 which is incompatible.
xarray 2025.3.1 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
treescope 0.1.9 requires numpy>=1.25.2, but you have numpy 1.23.5 which is incompatible.
blosc2 3.4.0 requires numpy>=1.26, but you have numpy 1.23.5 which is incompatible.
tf-keras 2.18.0 requires tensorflow<2.19,>=2.18, but you have tensorflow 2.17.1 which is incompatible.
ydf 0.12.0 requires protobuf<6.0.0,>=5.29.1, but you have protobuf 4.25.8 which is incompatible.
scikit-image 0.25.2 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
imbalanced-learn 0.13.0 requires numpy<3,>=1.24.3, but you have numpy 1.23.5 which is incompatible.
thinc 8.3.6 requires numpy<3.0.0,>=2.0.0, but you have numpy 1.23.5 which is incompatible.
Successfully installed gym-0.26.2 jax-0.4.30 jaxlib-0.4.30 ml-dtypes-0.4.1 numpy-1.23.5 protobuf-4.25.8 tensorboard-2.17.1 tensorflow-2.17.1
WARNING: The following packages were previously imported in this runtime:
  [google,gym,jax,jaxlib,ml_dtypes,numpy,tensorflow]
You must restart the runtime in order to use newly installed versions.
```

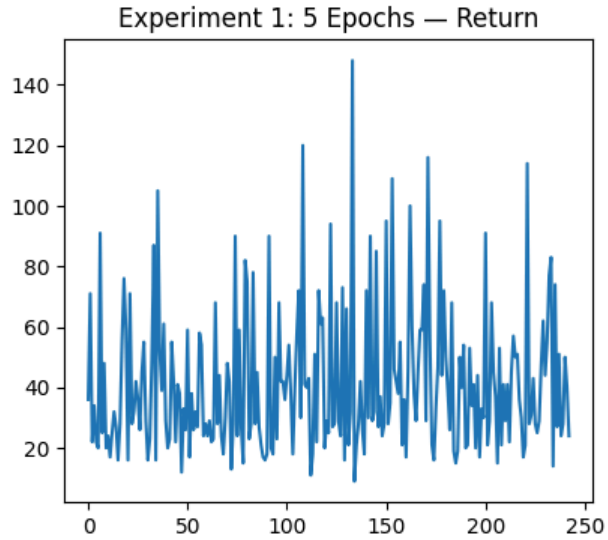
RESTART SESSION

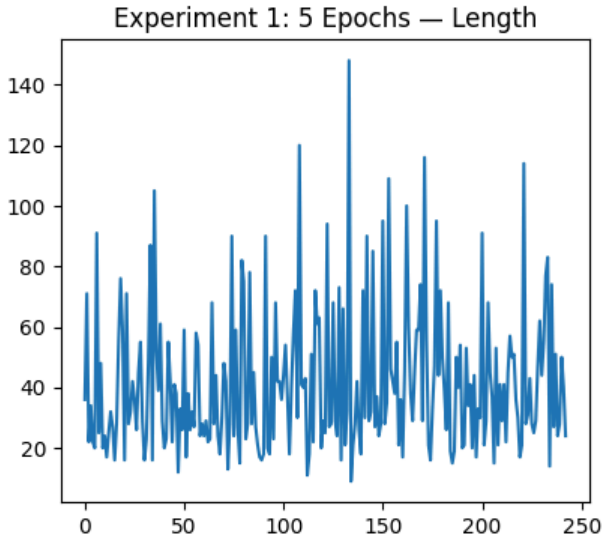
```
import gym
import numpy as np
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
import matplotlib.pyplot as plt

# PPO Agent
class PPOAgent:
    def __init__(self, obs_dim, act_dim, gamma=0.99, clip_ratio=0.2, hidden_sizes=(64, 64)):
        self.gamma = gamma
        self.clip_ratio = clip_ratio
        self.act_dim = act_dim

        inputs = keras.Input(shape=(obs_dim,))
        x = inputs
        for size in hidden_sizes:
            x = layers.Dense(size, activation='tanh')(x)
        logits = layers.Dense(act_dim)(x)
        values = layers.Dense(1)(x)

        self.model = keras.Model(inputs=inputs, outputs=[logits, values])
        self.optimizer = keras.optimizers.Adam(learning_rate=3e-4)

    def get_action(self, state):
        state = np.expand_dims(state, axis=0).astype(np.float32)
        logits, value = self.model(state)
        probs = tf.nn.softmax(logits)
        action = np.random.choice(self.act_dim, p=probs.numpy()[0])
        return action, probs[0, action].numpy(), value[0, 0].numpy()

    def compute_returns(self, rewards, dones, last_value):
        returns = []
        R = last_value
        for r, done in zip(reversed(rewards), reversed(dones)):
            R = r + self.gamma * R * (1. - float(done))
            returns.insert(0, R)
        return returns

    def train(self, states, actions, old_probs, returns, advantages, epochs=5):
        actions = np.array(actions)
        old_probs = np.array(old_probs)
        returns = np.array(returns)
        advantages = np.array(advantages)

        for _ in range(epochs):
            with tf.GradientTape() as tape:
                logits, values = self.model(states)
                values = tf.squeeze(values)
```

```
                probs = tf.nn.softmax(logits)
                indices = tf.range(tf.shape(logits)[0]) * self.act_dim + actions
                probs = tf.gather(tf.reshape(probs, [-1]), indices)

                ratio = probs / old_probs
                clipped = tf.clip_by_value(ratio, 1.0 - self.clip_ratio, 1.0 + self.clip_ratio)
                policy_loss = -tf.reduce_mean(tf.minimum(ratio * advantages, clipped * advantages))
                value_loss = tf.reduce_mean((returns - values) ** 2)
                loss = policy_loss + 0.5 * value_loss

            grads = tape.gradient(loss, self.model.trainable_variables)
            self.optimizer.apply_gradients(zip(grads, self.model.trainable_variables))



def run_training(env_name="CartPole-v1", epochs=5, hidden_sizes=(64, 64), clip_ratio=0.2):
    env = gym.make(env_name)
    obs_dim = env.observation_space.shape[0]
    act_dim = env.action_space.n
    agent = PPOAgent(obs_dim, act_dim, clip_ratio=clip_ratio, hidden_sizes=hidden_sizes)

    all_returns, all_lengths = [], []

    for epoch in range(epochs):
        states, actions, rewards, dones, probs, values = [], [], [], [], [], []
        obs = env.reset()
        if isinstance(obs, tuple): obs = obs[0]
        ep_len = ep_ret = 0

        while len(states) < 2048:
            action, prob, val = agent.get_action(obs)
            step_result = env.step(action)
            if len(step_result) == 5:
                next_obs, reward, terminated, truncated, _ = step_result
                done = terminated or truncated
            else:
                next_obs, reward, done, _ = step_result

            states.append(obs)
            actions.append(action)
            rewards.append(reward)
            dones.append(done)
            probs.append(prob)
            values.append(val)

            obs = next_obs
            ep_len += 1
            ep_ret += reward

            if done:
```

```
                    all_returns.append(ep_ret)
                    all_lengths.append(ep_len)
                    obs = env.reset()
                    if isinstance(obs, tuple): obs = obs[0]
                    ep_ret = ep_len = 0

            _, _, last_val = agent.get_action(obs)
            returns = agent.compute_returns(rewards, dones, last_val)
            advs = np.array(returns) - np.array(values)

            agent.train(np.array(states), actions, probs, returns, advs)
            print(f"Epoch {epoch+1} — Return: {np.mean(all_returns[-10:]):.2f} | Length: {np.mean(all_lengths[-10:]):.2f}")

        return all_returns, all_lengths



    def plot_results(returns, lengths, title):
        plt.figure(figsize=(10, 4))
        plt.subplot(1, 2, 1)
        plt.plot(returns)
        plt.title(f'{title} — Return')
        plt.subplot(1, 2, 2)
        plt.plot(lengths)
        plt.title(f'{title} — Length')
        plt.show()



    r1, l1 = run_training(epochs=5)
    plot_results(r1, l1, "Experiment 1: 5 Epochs")

    r2, l2 = run_training(epochs=5, hidden_sizes=(128, 128))
    plot_results(r2, l2, "Experiment 2: Hidden Size 128")

    r3, l3 = run_training(epochs=5, clip_ratio=0.4)
    plot_results(r3, l3, "Experiment 3: Clip Ratio 0.4")
```

```
Epoch 1 — Return: 32.80 | Length: 32.80
Epoch 2 — Return: 41.40 | Length: 41.40
Epoch 3 — Return: 45.40 | Length: 45.40
Epoch 4 — Return: 35.90 | Length: 35.90
Epoch 5 — Return: 41.50 | Length: 41.50
```






```
Epoch 1 — Return: 19.10 | Length: 19.10
Epoch 2 — Return: 26.70 | Length: 26.70
Epoch 3 — Return: 36.80 | Length: 36.80
Epoch 4 — Return: 37.00 | Length: 37.00
```